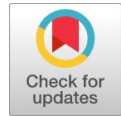# Performance Evaluation of Naive Bayes Classifier with and without Filter Based Feature Selection

D. Prabha, R. Siva Subramanian, S.Balakrishnan, M.Karpagam

Abstract: Customer Relationship Ma agement tends to analyze datasets to find insights about data which in turn helps to frame the business strategy for improvement of enterprises. Analyzing data in CRM requires high intensive models. Machine Learning (ML) algorithms help in analyzing such large dimensional datasets. In most real time datasets, the strong independence assumption of Naive Bayes (NB) between the attributes are violated and due to other various drawbacks in datasets like irrelevant data, partially irrelevant data and redundant data, it leads to poor performance of prediction. Feature selection is a preprocessing method applied, to enhance the predication of the NB model. Further, empirical experiments are conducted based on NB with Feature selection and NB without feature selection. In this paper, a empirical study of attribute selection is experimented for five dissimilar filter based feature selection such as Relief-F, Pearson correlation (PCC), Symmetrical Uncertainty (SU), Gain Ratio (GR) and Information Gain (IG).

Keywords: CRM, Prediction, Machine Learning, Naïve Bayes, feature selection

I. INTRODUCTION

CRM- business practice which helps to analyze potential customers of the organization and acquisition of new customers based on various techniques and strategies [1]. Customer retention and acquisition are the main objectives of every enterprise which helps to improve the business. Improper utilizing of CRM leads to failure of an enterprise [2]. The business goal is to make customer satisfy with their enterprises.

Customer acquisition is done by conducting various marketing strategies and customer profiles are collected. Through various forms of medium, the customer data are collected like customer feedback form, social media, purchases, email sign-ups, web sites, tele caller and in person collection [3].

Analyzing with these collected data, it helps us to find new customer and build the business to next level. Analyzing the dataset is curial one. Various machine learning algorithms and data mining algorithms helps to find out insights in the data [4].Machine Learning (ML) algorithms are efficient one in analyzing the datasets. Based upon the dataset characteristics, the ML algorithm can be split into three category supervised learning, reinforcement learning and unsupervised learning. In this work, supervised learning algorithm is studied and performed. Naive Bayes(NB) one of the classification model under supervised learning algorithm is applied and evaluated. NB a simple efficient model which predicates the feature with the given class label based upon Bayes theorem and make strong assumption independence between attributes and all attributes are equal[5]. Consider the dataset D $=\{y_1, y_2, ..., y_n \mid x_n\}$, in this attribute y_1, y_2 are independent and equal. In many real dataset the primary assumption of NB model is always violated and this approach leads to poor performance in model evaluation. With redundant, noisy, irrelevant and partially irrelevant data also performance of model degrades. This leads to increase of running time and with larger high dimensional datasets intractability and dimensionality problem also arise. To rectify these drawbacks and to enhance the performance of model, feature selections are carried out before performing model evaluation.

II. LITERATURE SURVEY

Feature(attribute or variable) selection is prominent one in supervised learning due to various causes such as complexity in running time, over fitting, generalization performance and interpretational problems [6]. In machine learning variable or feature or attribute selection is an crucial phase which leads to high performance of the classifier. The method of identifying optimal N subset attributes in the overall M features defines feature selection. Performing feature selection on datasets will improve accuracy of the model, reduces over fitting and reduce complexity time. Advancement in internet technology, mode of collection of datasets is exponentially increasing and datasets collected may contain irrelevant, partially relevant, redundant and noisy data which brings poor performance in the model evaluation. Primary weakness of Naive Bayes is about strong independence between features which are always violated in real time datasets. To rectify these issues performing feature selection become mandatory. Consider a database D with N features $(x_1, x_2, ... x_n)$ and Y class label, finding the optimal subset feature is a challenge.

The algorithm for identifying optimal subset feature should eliminate the irrelevant, noisy, partially relevant and redundant data. Let consider the dataset which deals with destroy of an object. The object properties includes dimension of the object, where to crush (in person place or in firm place), color of the object. In this dataset, except the feature color of the object, remaining attribute will contribute in model evaluation and thus by removing such partially relevant or irrelevant attribute, classifier performance can be improvement. Likewise the noisy data and redundant data would not contribute any information about the features. Eliminating these features becomes a prominent steps in machine learning feature selection.

Filter selector and wrapper selector are used for feature selection. Filter is attribute based ranking methodology. It depends upon the data samples to choose optimal subset features which are independent of the model. It uses some statistical technique to rank the feature [7]. A cut-off value is set and features above the grade are taken for the classifier evaluation. For large features, filter method is best one in selecting relevant feature when compared to wrapper model [8].

The wrapper method is based on search algorithm and it finds the best optimal subset by computing all feasible combinations of attributes and chooses the best attributes from the best combination. It uses classification model to evaluate the importance of an attribute set. Attribute selected highly depends on the predictor model. In spite, of best attribute selection, the method is computationally cost with large datasets [41].

III. METHODOLOGY

Naive Bayes (NB) belongs to the class of supervised classification algorithm. The supervised algorithm defines the presence of class label for the associated input features. NB is probabilistic classifier which is depends upon Bayes theorem. NB defines strong assumption that each attribute in datasets are independent and equal.To overcome the primary weakness of NB, different approaches are applied to improve the Naive Bayes model evaluation. These approaches can be grouped into two types. In the first method, the variable independence assumption of NB is relaxed using confined structure learning. These are some of the proposed methodology used to improve the NB classifier. SNB ("Semi-Naive Bayesian") classifier is applied. Four medical dataset are used for the problem, in which two datasets results stands identical and two datasets results tends to increase. Chebyshev Theorem is used to combine the attributes [9]. Dependencies among the features are identified using BSEJ ("Backward Sequential Elimination and Joining") and FSSJ methodology and comparison is done between the two methods. Results shows BSEJ ("Backward Sequential Elimination and Joining") tends to improve [10].Tree Augmented NB is proposed which maintains robustness and computational with respect to Naive Bayes. Results of wrapper, C4.5 and Naive Bayes are compared [11].Extended version of Tree Augmented NB is SuperParent-TAN, in which greedy search and SuperParent are explored [12].A Lazy Naïve Bayes approach is applied [13].Restricting model selection, AODE("Aggregating One Dependence Estimators)

reduces computational complexity and this helps to low variance[14].

Second method is to choose optimal subset from original features with the goal of maximize the accuracy of prediction. This method is called feature selection. Wrapper based forward selection is applied to select optimal features[15]. C4.5 decision tree based classifier is applied to select the attributes, which in turn selected features are applied to Naive Bayes to increase the prediction[16].

In our study second method, feature selection is experimented. Five different filter based attribute or feature selection is carried out. Relief-F, Pearson correlation (PCC), Symmetrical Uncertainty (SU), Gain Ratio (GR) and Information Gain (IG) are evaluated. Each selector is based on feature ranking techniques. Based upon the ranking, k highest features are chosen and in turn applied to classifier to improve the model.

A. Pearson Correlation Coefficient (PCC)

PCC coined by Karl Pearson computes the correlation between two variables [17].PCC values ranges from (-1 to +1). The value -1 indicates negative correlation, 0 indicates no correlation and +1 indicates positive correlation [18], [19]. Let X_i be feature and Y be the class label, the PCC can be denotes as [20]

$$\rho\ x, y = \frac{cov(X,Y)}{\sigma_x \sigma_y} \quad (1)$$

cov- covariance, σ_x - standard deviation X ,

σ_y - standarad deviation Y

PCC is used for binary regression problem and classification problem.

ρ can denoted in terms of expectation and mean

$$cov(X, Y) = E[(x - \mu_x)(y - \mu_y)] \quad (2)$$

$$\rho\ x, y = \frac{E[(x-\mu_x)(y-\mu_y)]}{\sigma_x \sigma_y} \quad (3)$$

whereμ_x - mean X , μ_y - mean Y , E - expectation

ρ can be denoted as

$$\rho\ x, y = \frac{E[XY] - E[x][Y]}{\sqrt{E[X^2] - [E[X]]^2}\sqrt{E[Y^2] - [E[Y]]^2}} \quad (4)$$

Strength of positive correlation lies between low(.1 to.3), medium(.3to.5) and high(.5 to 1.0).Strength of negative correlation lies between low (-0.1 to-0 .3), medium (-0.3 to-0 .5) and high(-0.5 to -1.0).

B. Relief-F

Kira and Rendell developed relief feature selector algorithm 21] and it is based on feature estimation. Relief algorithm originally constructed for binary classification issues, handles with two class problem and deals with continuous and discrete data types. Kononenko et al. proposed Relief-F enhanced version of relief method which overcomes the drawbacks [22]. Consider the instance $Y = \{y_1, y_2, \ldots y_n\}$ with class label, algorithm compute two nearby neighbours: one neighbour from its similar class denoted as hit H, then second one in dissimilar class denoted as miss M.

Depending with the value of Y, M, H every feature quantity estimate are updated [23].The algorithm assigns some relevant mark to each individual single features and attributes are selected which are above user defined threshold.

C. Information Gain (IG)

Information gain is univariate filter based feature selector named by Quinlan 1987. Consider the feature x=$(x_1,.....x_n)$ and y be the class label. IG computes each feature (x) information gain with respect to y class label. IG is computed with respect to entropy for our class label [24].

IG is expressed as [25],

$$IG(X,Y) = H(X) - H(X|Y) \qquad (5)$$

IG information gain, x feature, y class label, H(X│Y) observing x, H Entropy
Entropy X computed by [26],

$$H(X) = \sum_i p(x_i) \log p(x_i) \qquad (6)$$

X computed after observing Y[27],

$$H(X|Y) = -\sum_j P(y_i) \sum_i P(x_i|y_i) \log p(x_i|y_j) \qquad (7)$$

From the attribute, choose subset with high information gain.

D. Symmetrical Uncertainty (SU)

Enhanced information gain is used to overcome the drawbacks of IG bias towards variable with more values. SU value ranges between (0, 1) [28, 29]. SU value 0 imply no relationship between the two attribute and value 1 indicates one feature predict completely the other. Drop the attribute with SU lesser than prescribed value (Threshold).

SU expressed as [30], [31],

$$SU(X, Y) = 2 \frac{IG(XY)}{H(X) + H(Y)} \qquad (8)$$

Gain is expressed as [32],[33],

$$Gain(X,Y) = H(X) - H(X|Y) \qquad (9)$$

Entropy of X

$$H(X) = \sum_i p(x_i) \log p(x_i) \qquad (10)$$

H(X│Y) is expressed as [34],

$$H(X|Y) = -\sum_j P(y_i) \sum_i P(x_i|y_i) \log p(x_i|y_j) \qquad (11)$$

E. Gain Ratio (GR)

GR is an improvement of IG method which helps in reducing bias [28]. Let D be data specimen with m classes and then essential information is to analyze the same. I(D) can be expressed as [35],

$$I(D) = -\sum_{i=1}^{n} p_i \log_2 p_i \qquad (12)$$

p_i = Probability of sample which exit to class c_i

Let feature A holds V values. Then d_{ij} no of samples belongs to class c_i in subgroup d_j.

d_j holds the sample in D that possess 'a_j' of A. Then entropy based splitting the subset by A, is expressed as [36],

$$E(A) = \sum_{i=1}^{n} I(D) \frac{d_{1i}+d_{2i}+\cdots+d_{mi}}{d} \qquad (13)$$

Then encoding information gained by A

$$Gain(A) = I(D) - E(A) \qquad (14)$$

then, $info(D)$ as [37],

$$split\ info(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \log_2 \left(\frac{|D_j|}{|D|}\right) \qquad (15)$$

Then, Gain ratio (A) is expressed as [27],[38]

$$Gain\ Ratio(A) = \frac{Gain(A)}{spit\ info(A)} \qquad (16)$$

Feature with larger information gain are selected.

IV. EXPERIMENTS AND RESULTS

The dataset used for this experiment is obtained for UCI repository which is linked to direct bank marketing campaign and output analyze whether the customers will use the services are not [39].
Dataset contain totally 45211 instances with 17 attributes. Dataset are spitted into 70:30, it denotes training 70% and 30% testing and both dataset are mutually exclusive. The empirical study is conducted in two different perspectives. One is evaluation of Naive Bayes using feature selection and other one is evaluation of Naive Bayes without seeing feature selection. The purpose of the study is made to show how classifier performance can be improved by applying feature selection a preprocessing technique. First Naive Bayes model is executed without applying feature selection. Then each individual feature selector is applied with Naive Bayes and comparison is carried out with results.
The performances of the classifiers are computed based on accuracy, sensitivity or recall, specificity, F- measure and precision.
Accuracy measures totally number of samples correctly predicted. Specificity computed by number of negative instance predicated correctly, Sensitivity also referred as recall, compute number of positive instance predicted correctly [40].

Accuracy can be expressed as

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \qquad (17)$$

Sensitivity can be calculated as

$$Sensitivity = \frac{TP}{TP+FN} \qquad (18)$$

Specificity can be expressed as

$$Specificity = \frac{TN}{TN+FP} \qquad (19)$$

F-measure can be calculated using

$$F\text{-}measure = \frac{2*PRECISION*RECALL}{PRECISION+RECALL} \qquad (20)$$

Precision calculated using

$$Precision = \frac{TP}{TP+FP} \qquad (21)$$

The performances of the classifiers Naïve Bayes (NB), NB with Gain Ratio (NB & GR), NB with Pearson Correlation Coefficient (NB & PCC), NB with Information Gain (NB & IG), NB with Relief-F (NB & RELIEFF), NB with Symmetrical Uncertainty (NB & SU) are computed and the decisions are given in Tables 1&2. Figure 1, 2 and 3 are the graphical representations for the values given in Table 1 and 2.

Table 1: Summary of Accuracy, Specificity and Sensitivity of Naive Bayes with feature selection and without feature selection

Features & Model	Accuracy	Specificity	Sensitivity
NB	87.96	0.92	0.523
NB & GR	89.006	0.96	0.309
NB & PCC	89.41	0.96	0.366
NB & IG	89.02	0.95	0.402
NB & RELIEFF	89.31	0.95	0.45
NB & SU	89.006	0.96	0.309

Table2: Summary of Precision, F-Measure and Recall

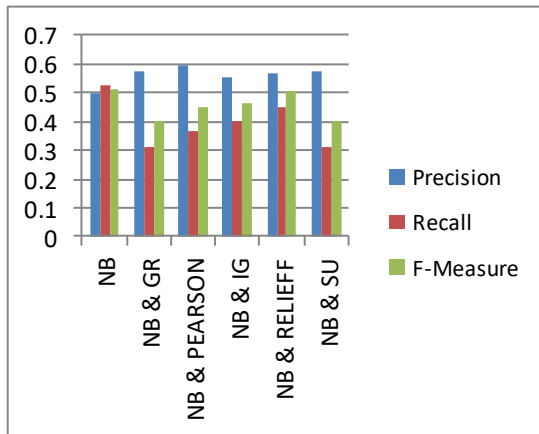| Features & Model	Precision	F-Measure	Recall
NB	0.496	0.509	0.523
NB & GR	0.574	0.401	0.309
NB & PCC	0.592	0.452	0.366
NB & IG	0.556	0.466	0.402
NB & RELIEFF	0.565	0.505	0.45
NB & SU	0.574	0.401	0.309

Fig. 1 Summary of Recall, Precision and F-measure
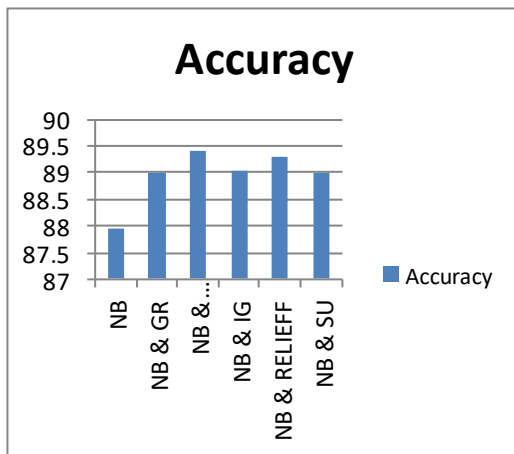

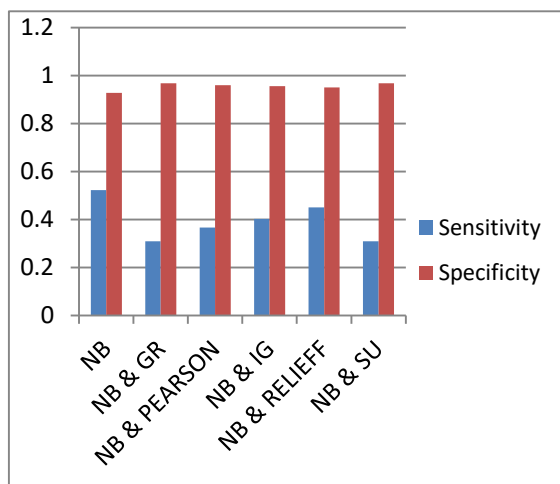**Fig. 2 Summary of Accuracy**

Fig. 3 Summary of Sensitivity & Specificity

V. CONCLUSION

The experiment study explores that there is better performance evaluation by using feature selection. Experiment from two different perspectives are conducted and studied. NB model evaluation without any feature selection technique is done and results are recorded. NB model evaluation after applying feature selection is carried out. Totally five filter supported feature ranking methods are carried out separately and results are noted. Results show that NB with attribute selection improve prediction well than NB without attribute selection. NB with Pearson Correlation Coefficient show greater accuracy with 89.41. As a future enhancement, ensemble method can be used to improve the performance. In Ensemble method, multiple selectors output can be combined as a single output and can be tried with still large data.

REFERENCES

1. Frederico Cruz-Jesus , Andreia Pinheiro, Tiago Oliveira, "Understanding CRM adoption stages: empirical analysis building on the TOE framework," Computers in Industry 109 (2019) 1–13
2. Zeynab Soltani et al., "The impact of the customer relationship management on the organization performance," Journal of High Technology Management Research 29 (2018) 237–24
3. 3   D. Prabha and R. S. Subramanian, "A survey on customer relationship management," 2017 (ICACCS), Coimbatore, 2017, pp. 1-5
4. Cristina-claudia osman, Ana -maria ghiran, "Extracting customer traces from CRMS: From Software to Process models," The 12 th International conference interdisciplinarity in engineering , Procedia manufacturing 32 (2019) 619-626.
5. Siva Subramanian Raju and Prabha Dhandayudam, "Prediction of customer behaviour analysis using classification algorithms," AIP Conference Proceedings 1952, 020098 (2018); doi: 10.1063/1.5032060
6. J.Weston et al., " Feature Selection for SVMs," In NIPS 13,2000
7. Mary Walowe Mwadulo, "A Review on Feature Selection Methods For Classification Tasks", IJCATR0506 Volume 5– Issue 6, 395 - 402, 2016
8. Sánchez-Maroño et al., " Filter Methods for Feature Selection – A Comparative Study," 178-187. 10.1007/978-3-540-77226-2_19.
9. Kononenko I. (1991), "Semi-naive bayesian classifier. In: Kodratoff Y. (eds) Machine Learning — EWSL-91,". EWSL 1991. Lecture Notes in Computer Science, vol 482. Springer, Berlin, Heidelberg
10. Pazzani M.J. (1996), "Searching for Dependencies in Bayesian Classifiers," Lecture Notes in Statistics, vol 112. Springer, New York, NY
11. Friedman, N Geiger, D. & Goldszmidt, "M, Machine Learning (1997) 29:131.https://doi.org/10.1023/A:10077465528199,"
12. E. Keogh, M. Pazzani, "Learning Augmented Bayesian Classifiers: A Comparison of Distribution-Based and Classification-Based Approaches," pp. 225-230, 1999
13. Zheng, Z. & Webb, G.I, "Lazy Learning of Bayesian Rules", Machine Learning, 41, 53-84 10.1023/A:1007613203719
14. Webb, G., Boughton, J. & Wang, Z, "Not So Naive Bayes: Aggregating One-Dependence Estimators," Machine Learning, 58 5-24. 10.1007/s10994-005-4258-6.
15. Pat langley,Stephanie sage, "Induction of Selective Bayesian Classifiers," UAI'94 Proceedings of the 10th international conference on Uncertainty in AI Pages 399-406.
16. C Ratanamahatana, D Gunopulos," Feature selection for the naive bayesian classifier using decision trees," Applied artificial intelligence, volume 17,Issue 5-6,pages 475-487
17. S. DeepaLakshmi and T. Velmurugan, "Empirical Study of Feature Selection Methods for High Dimensional Data," IJST, Vol 9(39), DOI: 10.17485/ijst/2016/v9i39/90599, October 2016

2157

18. Aytug Onam,Serder Korukoglu, "A feature selection model based on genetic rank aggregation for text sentinment classfication," Journal of Information science 1-14

19. Roberto Vega et al., "Assessment of feature selection and classification methods for recognizing motor imagery tasks from electroencephalographic signals," Artificial Intelligence Research, 2017, Vol. 6, No. 1

20. Ankita Mangal, Elizabeth A.Holm, "A Comparative Study of Feature Selection Methods for Stress Hotspot Classification in Materials," Integrating Materials and Manufacturing Innovation (2018) 7:87–95

21. Antonio Arauzo-Azofra, Jose Manuel Benitez,J.L.Casto, "A feature set measure based on Relief," Proc fifth Intl Conf,Artifical Intelligence,pp,279-305,1994

22. Sánchez-Maroño (2007) et. al, " Filter Methods for Feature Selection – A Comparative Study,"178-187. 10.1007/978-3-540-77226-2_19.

23. Salim Chikhi and Sadek Benhammada, " ReliefMSS: a variation on a feature ranking ReliefF algorithm," IJBIDM ,Volume 4,Issue 3/4,Nov 2009,Pages 375-390

24. 24.Jasmina NOVAKOVIĆ, Perica STRBAC, Dusan BULATOVIĆ, "Toward Optimal Feature Selection Using Ranking Methods And Classification Algorithms," Yugoslav Journal of Operations Research 21 (2011), Number 1, 119-135.

25. Thu ZarPhyu, NyeinNyeinOo, "Performance Comparison of Feature Selection Methods," MATEC Web of Conferences matec 42, 06002 (2016)

26. Jianli Ding, Liyang Fu, "A Hybrid Feature Selection Algorithm Based on Information Gain and Sequential Forward Floating Search," Journal of Intelligent Computing ,Volume 9 ,Number 3 September 2018

27. R.Porkodi, "Comparison of Filter Based Feature Selection Algorithms: An Overview," (IJIRTS) VOLUME 2, NUMBER2

28. Dr. Kadhim B.S. Al Janabi, Rusul Kadhim, "Data Reduction Techniques: A Comparative Study for Attribute Selection Methods," IJACST. ISSN 2249-3123 Vol 8, No 1 (2018), pp. 1-13

29. Venkatachalam K et al., "Effective Feature Set Selection And Centroid Classifier Algorithm For Web Services Discovery," International Journal of Pure and Applied Mathematics, Volume 119, No. 12, 2018, pp.1157-1172.

30. K Mani, P Kalpana, "A Review on Filter Based Feature Selection," IJIRCCE, Vol. 4, Issue 5, May 2016

31. Minghao Piao , Yongjun Piao and Jong Yun Lee, "Symmetrical Uncertainty-Based Feature Subset Generation and Ensemble Learning for Electricity Customer Classification," Symmetry 2019, 11, 498; doi:10.3390/sym11040498

32. D. SheelaJeyarani, A. Pethalakshmi, "An Efficient Fuzzy Entropy Based Feature Selection Algorithm for High Dimensional Data," International Journal of Advanced Research in Computer Science Volume 6, No. 6, June - July 2015

33. C. Sunil Kumar and R.J. Rama Sree, " Application Of Ranking Based Attribute Selection Filters To Perform Automated Evaluation Of Descriptive Answers Through Sequential Minimal Optimization Models," Ictact Journal On Soft Computing. 05.860-868.10.21917/ijsc.2014.0122

34. S.Senthamarai, N.Ramaraj, " A novel hybrid feature selection via Symmetrical Uncertainty ranking based local memetic search algorithm," Knowledge-Based Systems 23 (2010) 580–585

35. Suban Ravichrandan, Chandrasekaran Ramasamy, "Performance Comparison of Dimensionality Reduction Methods using MCDR," IJIRAE ISSN: 2349-2763 Issue 07, Volume 3 (July 2016)

36. N.S.Nithya and K Duraiswamy, "Gain ratio based fuzzy weighted association rule mining classifier for medical diagnostic interface," Vol. 39, Part 1, Feb 2014, pp. 39–52. c Indian Academy of Sciences

37. Anuj Sharma, Shubamoy Dey, " Performance Investigation of Feature Selection Methods and Sentiment Lexicons for Sentiment Analysis," International Journal of Computer Applications (0975 – 8887).

38. R. Praveena Priyadarsini , M.L.Valarmathi and S. Sivakumari, " Gain Ratio Based Feature Selection Method For Privacy Preservation," Ictact Journal On Soft Computing, April 2011, Volume: 01, Issue: 04

39. [Moro et al., 2014] S. Moro, P. Cortez and P. Rita, 'A Data-Driven Approach to Predict the Success of Bank Telemarketing," Decision Support Systems, Elsevier, 62:22-31, June 2014

40. Dhandayudam Prabha Krishnamurthi, Ilango, "Customer Behavior Analysis Using Rough Set Approach," Journal of Theoretical and Applied Electronic Commerce Research, vol. 8/ issuse 2/August, 2013, pp. 21-33

41. Yvan Saeys, Inaki Inza, Pedro Larranaga, "A review of feature selection techniques in bioinformatics," Bioinformatics ,Vol 23, Iss 19, 1 Oct 2007,Pg 2507-2517.

42. Balakrishnan S, K.Aravind, A. Jebaraj Ratnakumar, "A Novel Approach for Tumor Image Set Classification Based On Multi-Manifold Deep Metric Learning," International Journal of Pure and Applied Mathematics, Vol. 119, No. 10c, 2018, pp. 553-562.

43. A. Jebaraj Rathnakumar, S.Balakrishnan, "Machine Learning based Grape Leaf Disease Detection", JARDCS, Vol. 10, 775-780.

44. S. Vasu et al., "A Machine Learning Based Approach for Computer Security", JARDCS ,Vol.10, 915- 919.

AUTHORS PROFILE

Dr. D. Prabha is an Associate Professor in Department of Computer Science and Engineering at Sri Krishna College of Engineering and Technology, Coimbatore, Tamilnadu, India. She has completed her Under Graduate Studies and Post Graduate Studies in Computer Science and Engineering. She has received her Ph.D in Information and Communication Engineering from Anna University, Chennai in the year 2014. She is presently guiding 8 Ph.D. research scholars registered under Anna University, Chennai. Her research interests include Data Mining, Big Data Analytics and Soft Computing. She has published several papers in premier indexed journals.

Mr.R.Siva Subramanian is currently working as an Assistant Professor in Jawahar Engineering College, Chennai He is also pursuing his part time Ph.D in Anna University, Chennai. He has completed B.Tech IT in Srinivasa Institute of Engineering and Technology, Chennai. He has obtained M.Tech IT from Bharath University, Chennai. His research domain includes Data Mining, Customer Relationship Management and Big Data

Dr.S.Balakrishnan is a Professor in Department of Computer Science and Engineering at Sri Krishna College of Engineering and Technology, Coimbatore, Tamilnadu, India.

Dr.M.Karpagam has completed her Ph.D from Anna University, Chenni in April 2015. She is currently working as Professor in the department of Electronics and Communication Engineering, Sri Krishna College of Engineering and Technology, Coimbatore, India. Her areas of interest include Wireless Adhoc Sensor Networks and Image Processing.